

Lista 2

1) **Custos de manutenção**- O custo de manutenção de um determinado tipo de trator parece aumentar com a idade. O arquivo tractor.csv contém idades (anos) e custos de manutenção semestrais para $n = 17$ desses tratores.

(a) Crie um gráfico do custo de manutenção do trator em relação à idade.

(b) Encontre as estimativas de mínimos quadrados adequados ao modelo

$$\text{cost}_i = b_0 + b_1 \text{age}_i + e_i$$

de duas maneiras: primeiro usando o comando lm no R e, segundo, calculando uma correlação e desvios padrão [verificar se as respostas são idênticas]. Adicione a linha ajustada ao gráfico de dispersão.

(c) Suponha que você esteja pensando em comprar um trator de três anos, o que você esperaria sobre seus custos de manutenção semestral? Qual é o intervalo preditivo (custo) de 95% para o manutenção semestral do seu trator? Compare os pontos finais do intervalo com os valores observados de custo. O que você conclui sobre sua previsão a partir disso? Por que ou por que não é essa conclusão surpreendente?

2) Broadway

Sejam X e Y denotando os relatórios semanais sobre as vendas de ingressos de bilheteria para peças na Broadway em New York por duas semanas consecutivas, respectivamente, em outubro de 2004. (Você pode realmente fazer download de semelhantes dados de www.playbill.com.) A saída de regressão para este conjunto de dados é mostrada na tabela abaixo:

Variable	Coefficient	s.e.	t-value	p-value
Intercept	6805	9929	0.685	0.503
X	0.9821	0.01443	68.071	$< 2 \times 10^{-16}$
$n = 18 \quad R^2 = 0.9966 \quad s_e = 18007.56$				

Suponha que o modelo satisfaça as suposições usuais do modelo de regressão linear simples, e que o SST para Y seja $1,507773 \times 10^{12}$.

(a) Quais foram os graus de liberdade usados no cálculo de s_e ? Quais são SSE e SSR?

(b) Calcule a variância da amostra para Y (s_y^2) e correlação de amostra entre X e Y (r_{XY}).

(c) Suponha que a venda de ingressos na primeira semana para uma peça específica foi de \$822 000. O que é vendas esperadas para a mesma peça na semana seguinte?

(d) Suponha ainda que $\bar{X} = 822186,6$ e, o desvio padrão de X é $s_X = 302724,5$. Construir o intervalo de previsão de 95% para a estimativa em (c).

(e) Construir o intervalo de confiança de 95% para a inclinação da linha de regressão verdadeira β_1

(f) Algumas peças da Broadway usam a regra de que os resultados brutos de bilheteria da próxima semana serão os o mesmo desta semana. Isso é razoável? (Justifique / Refute usando um teste de hipótese apropriado.)

(g) Se Y e X fossem invertidos na regressão acima, o que você esperaria R^2 ser?